

## **Caracterização de Populações de Coqueiro-gigante-do-Brasil por Meio de Marcadores Microssatélites**



ISSN 1678-1961

Dezembro, 2013

*Empresa Brasileira de Pesquisa Agropecuária  
Embrapa Tabuleiros Costeiros  
Ministério da Agricultura, Pecuária e Abastecimento*

# ***Boletim de Pesquisa e Desenvolvimento 75***

## **Caracterização de Populações de Coqueiro-gigante-do-Brasil por Meio de Marcadores Microssatélites**

Francisco Elias Ribeiro  
Emiliano Fernandes Nassau Costa

Aracaju, SE  
2013

## **Embrapa Tabuleiros Costeiros**

Av. Beira Mar, 3250  
49025-040 Aracaju, SE  
Fone: (79) 4009-1344  
Fax: (79) 4009-1399  
www.cpatc.embrapa.br  
cpatc.sac@cpatc.embrapa.br

## **Comitê Local de Publicações da Embrapa Tabuleiros Costeiros**

Presidente: *Ronaldo Souza Resende*

Secretária-executiva: *Raquel Fernandes de Araújo Rodrigues*

Membros: *Ana Veruska Cruz da Silva, Edson Patto Pacheco, Élio César Guzzo, Hymerson Costa Azevedo, Joézio Luis dos Anjos, Josué Francisco da Silva Junior, Paulo César Falanghe Carneiro, Semíramis Rabelo Ramalho Ramos e Viviane Talamini*

Supervisão editorial: *Raquel Fernandes de Araújo Rodrigues*

Normalização bibliográfica: *Josete Cunha Melo*

Tratamento das ilustrações: *Ailla Freire de Azevedo*

Foto da capa: *Francisco Elías Ribeiro*

Editoração eletrônica: *Ailla Freire de Azevedo*

**1ª Edição (2013)**

On line (2013)

### **Todos os direitos reservados.**

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei no 9.610).

### **Dados Internacionais de Catalogação na Publicação (CIP)**

**Embrapa Tabuleiros Costeiros**

---

Ribeiro, Francisco Elías

Caracterização de populações de coqueiro-gigante-do-Brasil por meio de marcadores microsatélites / Francisco Elías Ribeiro, Emiliano Fernandes Nassau Costa. – Aracaju : Embrapa Tabuleiros Costeiros, 2013.

43 p. (Boletim de Pesquisa / Embrapa Tabuleiros Costeiros, ISSN 1678-1961; 75).

Disponível em: [http://www.cpatc.embrapa.br/publicacoes\\_2013/bp\\_75.pdf](http://www.cpatc.embrapa.br/publicacoes_2013/bp_75.pdf)

1. Coco. 2. Coqueiro-gigante. 3. População. 4. Genética de planta. I. Costa, Emiliano Fernandes Nassau. II. Título. III. Série.

CDD 634.61

# Sumário

<b>Resumo .....</b>	<b>04</b>
<b>Abstract.....</b>	<b>06</b>
<b>Introdução.....</b>	<b>07</b>
<b>Material e Métodos.....</b>	<b>08</b>
<b>Resultados e Discussão.....</b>	<b>15</b>
<b>Considerações Finais .....</b>	<b>36</b>
<b>Conclusões.....</b>	<b>37</b>
<b>Referências .....</b>	<b>38</b>

# Caracterização de Populações de Coqueiro-gigante-do-Brasil por Meio de Marcadores Microssatélites

Francisco Elias Ribeiro<sup>1</sup>

Emiliano Fernandes Nassau Costa<sup>2</sup>

## Resumo

O coqueiro é constituído de uma só espécie (*Cocos nucifera* L.) e de duas variedades principais, a gigante (*Typica*) e a anã (*Nana*). É a palmeira mais amplamente distribuída e extensivamente cultivada, além de ser uma das mais importantes espécies tropicais utilizadas pelo homem. Marcadores baseados em DNA têm sido indicados como os mais adequados nos estudos de diversidade genética em coco e os marcadores microssatélites (SSR) têm se mostrado como ferramenta poderosa em análises de estrutura de populações. O objetivo do presente trabalho foi avaliar a diversidade genética em dez populações de coqueiro-gigante-do-Brasil por meio de 13 marcadores microssatélites (SSR). A diversidade genética medida em 195 indivíduos de dez populações de coqueiro gigante do Brasil produziu um total de 68 alelos, com média de 5,23 alelos por loco. A diversidade de Nei (D) média foi de 0,459 e a heterozigiosidade observada média de 0,443. A diferenciação genética entre populações estimada foi  $F_{ST} = 0,1600$ , com intervalo de confiança (95%) de 0,122 a 0,199. A taxa de cruzamento aparente estimada foi de  $t_a = 0,92\%$  confirmando que a espécie se comporta como preferencialmente alógama. As estimativas de distâncias genéticas de Nei entre as populações variaram de 0,034 a 0,390. Com base nestas distâncias e no dendrograma correspondente propôs-se a formação de dois grupos, o primeiro formado pelas populações de Baía Formosa, Georgino

<sup>1</sup> Engenheiro-agrônomo, doutor em Genética e Melhoramento das Plantas, pesquisador da Embrapa Tabuleiros Costeiros, Aracaju, SE, [elias.ribeiro@embrapa.br](mailto:elias.ribeiro@embrapa.br).

<sup>2</sup> Agrônomo, doutor em Genética e Melhoramento das Plantas, pesquisador da Embrapa Tabuleiros Costeiros, Aracaju, SE, [emiliano.costa@embrapa.br](mailto:emiliano.costa@embrapa.br).

Avelino e São José do Mipibu e o segundo pelas populações de Japoatã, Pacatuba e Praia do Forte, com as demais populações individualizadas. A correlação matricial entre as distâncias genéticas de Nei e as distâncias geográficas foi positiva ( $r = 0,598$ ) e significativa ( $p = 0,0027$ ) pelo teste de Mantel. Este resultado sugere uma estruturação espacial da variabilidade genética entre as populações.

Palavras-chave: diversidade genética, SSR, estrutura populacional, marcadores moleculares.

# Characterization of Populations of Brazilian Tall Coconut by Microsatellite Markers

---

## Abstract

*The coconut palm (*Cocos nucifera* L.) is formed by a single species and two main varieties, the Tall (*Typica*) and Dwarf (*Nana*) coconut. This is the most spread and extensively cultivated palm and is one of the most important tropical species cultivated by man. DNA based markers have been indicated as the most appropriated for studies of the genetics of coconut. It has been shown that microsatellite (SSR) markers are important tools for the analysis of population structure. The aim of this paper was to evaluate the genetic diversity of ten Tall coconut populations in Brazil, through 13 microsatellite markers. The genetic diversity measured in 195 individuals of these populations produced a total of 68 alleles, with an average of 5.23 alleles per locus. Nei's genetic diversity (*D*) was 0.459 and the average observed heterozygosity equal to 0.443. The estimate of the genetic differentiation among populations was  $F_{ST} = 0.160$  with a 95% confidence interval of 0.122 to 0.199. The apparent outcrossing rate was  $t_a = 0.92$ , confirming that the species reproduces preferentially as allogamous. Estimates of Nei's genetic distance among populations varied from 0.034 to 0.390. Based on these distances and the corresponding dendrogram, it is proposed that two groups of populations can be created, the first one containing populations from Baía Formosa, Georgino Avelino and São José do Mipibu and the second group by populations from Japoatã, Pacatuba and Praia do Forte. The remaining populations were individualized. The correlation matrix between Nei's genetic distances and corresponding geographic distances was positive ( $r = 0.598$ ) and significant at 1% probability ( $P = 0.0027$ ).*

*Index terms: genetic diversity, SSR, population structure, molecular markers.*

## Introdução

O coqueiro é constituído de uma só espécie (*Cocos nucifera* L.) e composto por duas variedades principais: a gigante (*Typica*) e a anã (*Nana*). É a palmeira mais amplamente distribuída e extensivamente cultivada, além de ser uma das mais importantes espécies tropicais utilizadas pelo homem (PERSLEY, 1992).

O sudeste asiático é o provável centro de origem da espécie, pois é nessa região que ocorre a maior variabilidade morfológica, maior número de nomes locais, diferentes usos da planta e maior número de espécies de insetos associados (PERSLEY, 1992). Do sudeste asiático o coqueiro foi levado para a Índia e daí para o leste africano. Após o descobrimento do Cabo da Boa Esperança, essa planta foi levada para o oeste africano e, desta região para as Américas e toda a região tropical do Globo (PURSEGLOVE, 1975). A variedade gigante foi introduzida no Brasil em 1553, proveniente da ilha de Cabo Verde.

A variedade gigante (*Typica*) é de porte alto, produz um número médio de frutos grandes, é tardio, iniciando o florescimento entre 5 e 7 anos, e sua reprodução é predominantemente por fecundação cruzada. Por outro lado, a variedade anã (*Nana*) apresenta menor porte, produz muitos frutos de tamanho pequeno, é precoce (inicia a produção entre 2 e 3 anos de idade) e se reproduz predominantemente por autofecundação (SIQUEIRA et al., 1994).

No Nordeste, onde se concentra a maior parte dos coqueirais brasileiros, há algumas populações de coqueiro-gigante que foram implantadas há mais de 80 anos e que se conhece muito pouco sobre a sua variabilidade genética. A obtenção dessas informações é fundamental para a utilização em programas de melhoramento da cultura, pois essas populações podem se constituir em uma excelente fonte de germoplasma adaptado (RIBEIRO, 1993).

O coqueiro-gigante se desenvolveu no Brasil há mais de 450 anos. Atualmente estima-se a existência de uma área plantada de aproximadamente 274.000 hectares, com produção em torno de 2,8 milhões de toneladas de frutos por ano (FAO, 2012), distribuída desde a linha do Equador até o Trópico de Capricórnio, e situada ao longo da faixa litorânea, que se estende do Pará ao Rio de Janeiro. Todavia a maioria desses coqueirais está localizada no Nordeste (RIBEIRO et al., 2000). As populações de coqueiro assim distribuídas foram se adaptando às diferentes condições de ambiente, tornando-as divergentes entre si (RIBEIRO et



al., 1999) ou por meio de introduções independentes e se caracterizando como ecótipos da variedade gigante.

Marcadores baseados em DNA têm sido indicados como os mais adequados nos estudos de diversidade genética em coco (LEBRUN et al., 1995; WADT et al., 1997), enquanto que isoenzimas (BENOIT; GHESQUIÈRE, 1984) e polifenóis de folha (JAY et al., 1989) não proporcionam resultados conclusivos. Estudos de diversidade realizados com marcadores RAPD (WADT et al., 1997) propiciaram o conhecimento da diversidade genética em algumas populações de coqueiro-gigante e anão. No entanto, marcadores microsatélites (SSR) têm se mostrado como a ferramenta mais poderosa em análises de estrutura de populações, devido às características deste marcador: multi-alelico, altamente polimórfico e baseado na PCR (CHASE et al., 1996; MORGANTE et al., 1996). Diante do avanço das técnicas de biologia molecular e das vantagens oferecidas pelos marcadores microsatélites, principalmente por serem codominantes, baseados na reação automatizada de PCR e de menor custo, optou-se por utilizar esta classe de marcadores para as análises que se propõem neste trabalho.

O objetivo do presente trabalho foi avaliar a diversidade genética em dez populações de coqueiro-gigante-do-Brasil por meio de marcadores microsatélites (SSR) visando à caracterização da variabilidade genética.

## **Material e Métodos**

### **Material vegetal**

Na estratégia de prospecção genética e coleta do coqueiro-gigante, é importante definir as populações a serem investigadas. A seleção das populações para este estudo foi feita após a prospecção genética de coqueiro-gigante-do-Brasil e definidas como populações puras (típicas) da variedade gigante. A pureza é avaliada com base nos critérios de legitimidade, homogeneidade e isolamento. A primeira é definida em função da idade, sendo o ideal populações com mais de 80 anos, pois como a variedade anã foi introduzida no país em 1925, evitar-se-ia o risco da ocorrência de híbridos naturais entre as duas variedades. Já a homogeneidade leva em consideração a constituição das populações, onde estas devem ser compostas apenas por plantas da variedade gigante. As condições de isolamento dessas populações são outro fator importante, e para as quais deve haver uma distância mínima de 1000 metros, ou pelo menos 500 metros se

houver barreira de vegetação isolando-as de outros coqueirais, principalmente de anões (RIBEIRO et al., 2002).

Foram identificadas, por meio de prospecções, como legítimas e homogêneas da variedade gigante e que atendem às condições de isolamento, no Brasil, as seguintes populações: Praia do Forte (BA), Pacatuba e Japoatã (SE), Merepe e Santa Rita (PE), São José do Mipibu, Baía Formosa e Georgino Avelino (RN), Luís Correia (PI) e Barreirinhas (MA). Essas dez populações foram estudadas no presente trabalho (Figura 1), nas quais as áreas de coletas estão localizadas, de acordo com as coordenadas geográficas, altitude e número de plantas coletadas por população (Tabela 1).



**Figura 1.** Mapa de localização das populações de coqueiro-gigante do Brasil (pontos vermelhos/azul), coletadas na região Nordeste. Aracaju, SE, 2013.

**Tabela 1.** Populações de coqueiro-gigante-do-Brasil coletadas, com o número de plantas avaliadas, procedência, coordenadas geográficas e altitudes. Aracaju, SE, 2013.

População	Número de plantas	Procedência (Estado)	Latitude (S)	Longitude (W. Gr.)	Altitude (m)
Santo Inácio (SI)	17	Maranhão	02° 34' 54"	42° 45' 15"	2
Luís Correia (LC)	17	Piauí	02° 57' 59"	41° 35' 45"	7
Baía Formosa (BF)	30	Rio Gr. do Norte	06° 23' 14"	35° 03' 31"	9
Georgino Avelino (GA)	21	Rio Gr. do Norte	06° 10' 57"	35° 06' 02"	2
S. J. Mipibu (SJM)	18	Rio Gr. do Norte	06° 06' 43"	35° 14' 36"	8
Santa Rita (SR)	21	Pernambuco	07° 10' 54"	34° 53' 01"	11
Merepe (Me)	20	Pernambuco	08° 28' 38"	34° 59' 52"	2
Japoatã (Jp)	18	Sergipe	10° 27' 08"	36° 42' 42"	232
Pacatuba (Pc)	19	Sergipe	10° 30' 29"	36° 35' 37"	6
Praia do Forte (PF)	14	Bahia	12° 34' 09"	37° 59' 48"	2

A amostragem foi feita levando-se em consideração o tamanho das populações e o estado fitossanitário das plantas. Para composição das amostras, foram coletados folíolos de 40 plantas por população, onde de cada planta, foi retirado um segmento de aproximadamente 50 cm de folíolos da folha mais jovem (flecha), que foram acondicionados em sacos plásticos individuais, com o número da planta, data e nome da população e armazenados em isopor com gelo. Em seguida as amostras foram encaminhadas para a Embrapa Tabuleiros Costeiros, em Aracaju - Sergipe, onde foram liofilizadas e armazenadas em freezer à -20 °C para subsequente extração de DNA e análise de microsatélites (SSR).

## Métodos laboratoriais

As amostras de tecido foliar de 400 plantas coletadas foram liofilizadas na Embrapa Tabuleiros Costeiros, em Aracaju (SE) e transportadas para o Laboratório de Biotecnologia da Embrapa Arroz e Feijão em Goiânia, Goiás, onde foram feitas extração e quantificação de DNA. Depois, tanto o DNA extraído e quantificado, quanto amostras de tecido foliar, foram levadas para o CIRAD (Centre de Coopération Internationale em Recherche Agronomique pour le Développement), em Montpellier, França. Como a maioria do DNA extraído no Brasil não amplificou, foram feitas novas extrações de DNA, quantificação, amplificação e análise de SSR.

O DNA foi extraído de acordo com protocolo CTAB modificado e adaptado para o coqueiro, denominado de extração de DNA genômico Par Quiagen, utilizando tampão de extração MATAB, desenvolvido e utilizado pelo CIRAD (LEBRUN et al., 1998; BAUDOUIN; LEBRUN, 2002). O tampão de extração MATAB é composto da seguinte forma: Tris-Cl, pH=8,0, (100 mM); NaCl (1,4 M); EDTA (20 mM); MATAB (2%); PEG 6000 (1%); Sulfito de Sódio (0,5%) e o TE composto de Tris-Cl (50mM); NaCl (0,7M) e EDTA (10mM), ajustado para pH=7,0. A concentração do DNA na solução de trabalho foi de 2,5 ng/mL.

A concentração de DNA extraído foi quantificada por eletroforese em gel de agarose a 1,2% com tampão TBE 1X e também por quantificação automática por fluorometria. Para visualização do resultado da amplificação via PCR usou-se gel de agarose a 2%. A migração foi feita a 120 V por 60 minutos para a quantificação e 100 V durante 90 minutos para a reação de amplificação da PCR.

Para a reação de PCR foi preparado um volume final de 25  $\mu$ L, contendo uma mistura composta por 2,5  $\mu$ L de tampão PCR 10X; 2,0  $\mu$ L de dNTP (2 mM); 0,25  $\mu$ L de MgCl<sub>2</sub> (50 mM); 0,5  $\mu$ L de Primer Forward (10  $\mu$ M); 0,5  $\mu$ L de Primer Reverse (10  $\mu$ M); 0,5  $\mu$ L de Taq DNA Polimerase (2U/  $\mu$ L); 5  $\mu$ L de DNA genômico (2,5 ng/ $\mu$ L) e 13,75  $\mu$ L de água estéril. O ciclo de reação de PCR consistiu de uma desnaturação inicial a 94 °C durante cinco minutos, seguido de 36 ciclos com temperaturas de 94 °C por 30 segundos para desnaturação, um minuto a 51 °C para anelamento e um minuto a 72 °C para extensão, após isso, mais um período adicional de extensão por cinco minutos a 72 °C (LEBRUN et al., 1998; BAUDOUIN; LEBRUN, 2002).

Para a amplificação dos SSRs utilizaram-se 13 pares de primers específicos para o coqueiro (Tabela 2), desenvolvidos pelo CIRAD e a visualização das bandas foi feita em gel de poliacrilamida com marcação dos produtos amplificados por meio de compostos fluorescentes, em seqüenciador Licor IR2 4200. As bandas foram lidas e genotipadas de acordo com o comprimento do alelo, expresso em número de pares de base pela comparação com o marcador "ladder" padrão para o coqueiro (1 Kb).

**Tabela 2.** Comprimento observado dos alelos e temperatura de anelamento da PCR nos 13 locos de SSR avaliados em dez populações de coqueiro-gigante-do-Brasil. Aracaju, SE, 2013.

Loco SSR	Comprimento observado (pb)	Temperatura de anelamento (°C)
CnCir A3	228-240	51
CnCir A9	089-103	51
CnCir B6	196-208	51
CnCir B12	157-181	51
CnCir C7	157-167	51
CnCir C12	163-183	51
CnCir E2	115-165	51
CnCir E10	232-246	51
CnCir E12	164-174	51
CnCir F2	193-205	51
CnCir G11	188-210	51
CnCir H4'	218-230	51
CnCir H7	133-141	51

A análise de marcadores microsatélites foi feita com base em 13 locos SSR de boa qualidade, que foram desenvolvidos e selecionados pelo CIRAD conforme descrito por Baudouin e Lebrun (2002), sendo que a avaliação das populações foi feita com um total de 195 plantas. A temperatura de anelamento e o comprimento observado dos alelos encontram-se, na Tabela 2.

## Análises biométricas

### Estatísticas descritivas

Os indivíduos analisados foram genotipados e a matriz de genótipos foi utilizada para cálculos de frequências alélicas e distâncias genéticas entre indivíduos e populações. A partir da genotipagem de todos os indivíduos, foram obtidas as frequências gênicas (alélicas) e genotípicas em cada loco.

As estatísticas descritivas tais como: frequências alélicas, número total de alelos, número de alelos por loco, número de alelos por loco polimórfico, diversidade gênica ( $D$ ), heterozigosidade observada ( $H_o$ ), número de locos polimórficos, proporção de locos polimórficos e alelos exclusivos foram utilizadas na caracterização das populações.

Também foi utilizada a classificação dos indivíduos de acordo com a sua população de referência conforme o critério de Rannala e Mountain (1997), para auxiliar na caracterização das populações. Este método tem como objetivo geral associar os indivíduos a uma população, dentre o grupo de populações da amostra, denominadas de populações de referência, usando as informações dos marcadores SSR, com base nas frequências alélicas. O método consiste em responder a seguinte questão: entre as populações do grupo de referência, qual é a mais provável de ser a origem de um determinado indivíduo. Para isso é assumido equilíbrio de Hardy-Weinberg dentro de cada loco e equilíbrio de ligação entre os locos. Essa associação é feita com base nos escores das probabilidades. Quando um determinado indivíduo está associado a uma população com probabilidade superior a 95%, diz-se que esse indivíduo é bem classificado e é, portanto, considerado não ambíguo.

Essas estatísticas foram estimadas utilizando o programa GENECLASS2 (PIRY et al., 2004) utilizando-se a opção “associação dos indivíduos”, conforme o critério de Rannala e Mountain (1997), para alocar os indivíduos às populações de referência e a opção estatísticas para as demais estatísticas descritivas.

## Estrutura de variabilidade genética

A estruturação da variabilidade genética foi avaliada com base nas estimativas das estatísticas F de Wright (WRIGHT, 1951), obtidas conforme o método de Weir e Cockerham (1984). Estas estimativas foram obtidas por meio do programa GENETIX 4.03 (BELKHIR et al., 2001) utilizando a opção Fstats.

A estrutura genética das populações foi caracterizada pelos parâmetros: índice médio de fixação dentro das populações ( $f$ ), índice de fixação total das populações ( $F$ ) e divergência genética entre as populações ( $D_p$ ). Estimou-se o intervalo de confiança a 95% de probabilidade pelo método de reamostragem *bootstrap* sobre a média de todos os locos, utilizando 10.000 repetições por loco. Estes parâmetros correspondem, respectivamente aos parâmetros  $F_{IS}$ ,  $F_{IT}$  e  $F_{ST}$  de Wright.

A diversidade genética e as estatísticas F foram estimadas sob modelo aleatório de acordo com Weir (1996), em que as populações amostradas são consideradas como representativas do conjunto das populações e com uma história evolutiva comum.

O estimador  $\hat{F}_p$  pode ser entendido como uma medida de diferenciação entre subpopulações e é consequência da correlação entre genes de diferentes indivíduos na mesma subpopulação (COCKERHAM, 1973; WEIR; COCKERHAM, 1984). Já o parâmetro  $F$  é um coeficiente de endogamia de Wright que envolve todos os indivíduos que representam o conjunto de todas as subpopulações, ou seja, é um estimador de correlação entre genes dentro de indivíduos (COCKERHAM, 1969; COCKERHAM, 1973; VENCovsky, 1992).

## Sistema Reprodutivo

A taxa aparente de fecundação cruzada foi obtida após o teste de aderência às proporções de equilíbrio de Hardy-Weinberg, conforme definido por Weir (1996), utilizando o programa TFGA (Miller, 1997). O teste de aderência foi realizado utilizando-se o método convencional de Monte Carlo utilizando 10 *batches* com 1000 permutações por *batch*. Admitindo que as populações se encontram em equilíbrio de Wright, foi estimada a taxa aparente de fecundação cruzada conforme procedimento descrito por Vencovsky (1994), sendo  $t_a = (1-f)/(1+f)$ , onde:  $t_a$  = taxa aparente de cruzamento e  $f$  = índice de fixação intrapopulacional. O índice de fixação intrapopulacional ( $f$ ), que corresponde ao  $F_{IS}$  de Wright, foi estimado por meio do programa GENETIX 4.03 (BELKHIR et al., 2001) utilizando a opção Fstats.

## Padrão espacial da variabilidade genética

O padrão espacial da variabilidade genética foi avaliado a partir da estimativa do coeficiente de correlação de Pearson ( $r$ ) entre as matrizes de distâncias genéticas de Nei (Nei, 1972) e de distâncias geográficas entre as populações. A significância desta correlação foi testada pela estatística  $Z$  de Mantel (MANLY, 1986a; 1986b), utilizando 10.000 permutações aleatórias. As estimativas das distâncias de Nei foram obtidas utilizando-se o programa GENETIX 4.03 (BELKHIR et al., 2001).

Com o propósito de analisar os padrões de variação espacial em um contexto multivariado, foi obtida a estimativa do coeficiente de correlação de Pearson ( $r$ ) entre as matrizes de distâncias genéticas de Nei e as respectivas distâncias geográficas entre as populações. A significância dessa correlação matricial não pode ser testada por testes estatísticos usuais, por apresentar problemas de independência entre elementos nas matrizes, por isso utilizou-se a estatística  $Z$  de Mantel (1967). Estas análises foram feitas no programa GENETIX 4.03 (BELKHIR et al., 2001).

## Aspectos histórico-evolutivos

Para o entendimento dos aspectos histórico-evolutivos das populações, procurou-se representar o padrão de distribuição dessas populações em gráficos tridimensionais, utilizando-se uma técnica de análise multivariada denominada Análise Fatorial de Correspondência (AFC). O objetivo deste método é produzir uma representação geométrica das medidas de dissimilaridade entre unidades, apresenta ainda uma decomposição hierárquica da diversidade, que permite distinguir as tendências fundamentais das variações particulares (PERRIER et al., 1999). A AFC é um tipo de análise canônica particularmente adaptada para descrever as associações entre variáveis qualitativas. A técnica consiste na redução de um hiperespaço  $n$  dimensional a um espaço de poucas dimensões, neste caso, tridimensional, permitindo observar, neste espaço simplificado, a posição relativa dos objetos (populações) e sua possível associação em grupos de similaridade, sendo que os pontos mais próximos representam maior similaridade genética entre as populações. Para essa finalidade utilizou-se o programa GENETIX 4.03 (BELKHIR et al., 2001).

Utilizaram-se as distâncias genéticas de Nei para construção de dendrogramas não enraizados para avaliação do padrão de divergência genética entre as populações. Para a construção dos dendrogramas foi utilizado o método Neighbor Joining (SAITOU; NEI, 1987), por meio do programa PHYLIP 3.6 (FELSENSTEIN, 2004).

O método Neighbor Joining é um método de construção de árvores não enraizadas, que leva em conta a filogenia e não requer taxas evolutivas constantes. Esse algoritmo une unidades taxonômicas de modo a minimizar o comprimento total da árvore inferida, ou seja, busca uma árvore com a menor soma de comprimentos de ramos.

## Resultados e Discussão

### Caracterização da variabilidade genética

#### Estatísticas descritivas

A diversidade alélica medida nos 195 indivíduos das dez populações de coqueiro-gigante-do-Brasil produziu perfil de locos simples. O número total de



alelos, a diversidade genética (D), e a heterozigosidade observada ( $H_o$ ) para cada loco são apresentados, na Tabela 3. Todos os 13 locos em combinação produziram um total de 68 alelos, com média de 5,23 alelos por loco, variando de dois alelos para o loco CnCir E12 a 13 alelos para o loco CnCir E2. Não foram encontrados locos monomórficos para essas amostras no conjunto das populações.

**Tabela 3.** Resultados de 13 locos microsatélites em dez populações de coqueiro-gigante-do-Brasil, em relação ao número de alelos por loco, diversidade genética (D) e heterozigosidade observada ( $H_o$ ) nos locos. Aracaju, SE, 2013.

Número	Nome SSR	Nº de alelos	D	$H_o$
1	CnCir A3	3	0,034	0,036
2	CnCir A9	4	0,513	0,484
3	CnCir B6	5	0,613	0,531
4	CnCir B12	9	0,515	0,464
5	CnCir C7	5	0,495	0,470
6	CnCir C12	6	0,378	0,355
7	CnCir E2	13	0,711	0,671
8	CnCir E10	4	0,418	0,443
9	CnCir E12	2	0,300	0,302
10	CnCir F2	3	0,510	0,536
11	CnCir G11	7	0,638	0,640
12	CnCir H4'	3	0,330	0,311
13	CnCir H7	4	0,512	0,518
<b>Média</b>		<b>5,23</b>	<b>0,459</b>	<b>0,443</b>
<b>Total</b>		<b>68</b>		

O número médio de alelos por loco encontrado (5,2) foi semelhante a outros trabalhos realizados com coqueiro no mundo. Rivera et al. (1999), avaliando 20 plantas com 38 locos SSR, encontraram média de 5,2 alelos por loco, variando de 2 a 9 alelos por loco, com um total de 198 alelos. Já Perera et al. (2000), avaliando 75 plantas com 8 locos SSR, encontraram média de 6,3 alelos por loco, variando de 3 a 9 alelos por loco, com um total de 50 alelos (Tabela 3).

A diversidade genética (D) variou de 0,034 no loco CnCir A3 a 0,711 no loco CnCir E2, com média de 0,459. Estes locos também apresentaram a menor (0,036) e a maior (0,671) heterozigosidade observada, respectivamente, com média de 0,443, considerando todas as populações (Tabela 3).

O número de indivíduos amostrados por população (N), número de locos (L), número de locos polimórficos (NP), proporção de locos polimórficos (P), número total de alelos (A), número médio de alelos por loco (NA), número de alelos por loco polimórfico (ALP); diversidade genética (D) e heterozigosidade observada ( $H_o$ ) nas populações são apresentados, na Tabela 4. Considerando todos os locos, a diversidade genética, que corresponde a heterozigosidade esperada sob equilíbrio de Hardy-Weinberg, variou de 0,358 a 0,552 entre as populações, sendo o menor valor encontrado na população de Georgino Avelino (4) e o maior valor na população de Japoatã (8), com média geral de 0,459. Já a heterozigosidade observada variou entre as populações de 0,315 na população de Georgino Avelino (4) a 0,553 na população de Pacatuba (9), com média geral de 0,443.

**Tabela 4.** Estimativas de parâmetros genéticos de diversidade em 10 populações (Pop) de coqueiro-gigante-do-Brasil, avaliados por microsatélites, sendo N: número de indivíduos amostrados; L: número de locos microsatélites; NP: número de locos polimórficos; P: proporção de locos polimórficos; A: número total de alelos por população; NA: número médio de alelos por loco; ALP: número de alelos por loco polimórfico; D: diversidade genética e  $H_o$ : heterozigosidade observada. Aracaju, SE, 2013.

Pop	N	L	NP	P	A	NA	ALP	D	$H_o$
1	17	13	12	0,923	37	2,85	3,08	0,455	0,484
2	17	13	12	0,923	40	3,10	3,33	0,451	0,443
3	30	13	12	0,923	41	3,20	3,42	0,415	0,375
4	21	13	10	0,769	40	3,10	4,00	0,358	0,315
5	18	13	12	0,923	38	2,92	3,17	0,426	0,371
6	21	13	12	0,923	40	3,10	3,33	0,502	0,530
7	20	13	11	0,846	36	2,80	3,27	0,404	0,361
8	18	13	13	1,000	48	3,70	3,70	0,552	0,520
9	19	13	13	1,000	47	3,62	3,62	0,504	0,553
10	14	13	13	1,000	42	3,23	3,23	0,521	0,479
Média	19,5	13	12	0,923	40,9	3,16	3,42	0,459	0,443

Os menores valores para D e  $H_o$  foram encontrados na mesma população, Georgino Avelino (4), apresentando também a menor proporção de locos polimórficos (77%), enquanto que os maiores valores para estes parâmetros foram encontrados nas populações de Japoatã (8) e Pacatuba (9), respectivamente. Estas populações apresentaram 100% de locos polimórficos (Tabela 4).

A diversidade genética (D) estimada neste trabalho foi inferior à encontrada por Perera et al. (2001) em coleções *ex situ* de coqueiro-gigante no Sri Lanka. Esses autores encontraram valores que variaram de 0,426 a 0,846 nas populações, com média de 0,682.

A proporção de locos polimórficos (P) variou de 0,770 a 1,000 sendo o menor valor encontrado na população de Georgino Avelino (4) e o maior valor encontrado nas populações de Japoatã (8), Pacatuba (9) e Praia do Forte (10), sendo que essas populações situam-se ao Sul da região de coleta. A média global foi de 0,923, considerando todas as populações (Tabela 4).

Em relação à proporção de locos polimórficos, apenas as populações que formam o grupo do Sul (Japoatã, Pacatuba e Praia do Forte) apresentaram 100% de polimorfismo. Dentro do outro grupo, o de Natal, composto pelas populações de Baía Formosa, Georgino Avelino e São José do Mipibu, a população de Georgino Avelino diferiu das outras duas em relação a este parâmetro, com 77% de locos polimórficos, comparado com 92,3% de polimorfismo apresentado pelas outras duas populações.

Para os 13 locos utilizados, o número total de alelos por população variou de 36 na população de Merepe (7) a 48 na população de Japoatã (8), com média geral de 40,9 alelos por população e o número médio de alelos por loco variou de 2,8 a 3,7, sendo o maior valor encontrado na população de Japoatã e o menor na população de Merepe, com média geral de 3,16 alelos por loco. Vale ressaltar que as populações que apresentaram os maiores números de alelos totais, 48 (Japoatã), 47 (Pacatuba) e 42 (Praia do Forte), não foram as populações com maior tamanho de amostra (Tabela 4).

O número de indivíduos avaliados variou de 14 na população da Praia do Forte a 30 na população de baía Formosa, com média de 19,5 indivíduos por população. Já o número de locos polimórficos variou de 11 na população de Merepe a 13 nas populações que formam o grupo do Sul, com média de 12 locos polimórficos por população (Tabela 4).

O número de alelos exclusivos e as respectivas frequências por loco são apresentados na Tabela 5. O número de alelos exclusivos ou privados observados variou de 1 a 3, sendo um total de 12 alelos privados (17,65%) distribuídos em seis das dez populações avaliadas. Embora as populações compartilhem a maior parte dos 68 alelos, há um pequeno número de alelos que caracterizam

determinadas populações, sendo três nas populações de Georgino Avelino, Japoatã e Pacatuba e um nas populações de Santa Rita, Baía Formosa e Praia do Forte. Quatro desses alelos podem ser considerados como comuns localizados, ou seja, estão presentes em uma única população, porém com frequência  $\geq 5\%$ . Neste caso, para capturar esses alelos poder-se-ia utilizar a estratégia de coletar pequenas amostras em muitas populações e é nessa categoria que os melhoristas concentram seus esforços, uma vez que os alelos comuns dispersos não apresentam problemas, podendo ser capturados mesmo em pequenas amostras coletadas de poucas populações.

**Tabela 5.** Alelos exclusivos por loco e suas respectivas frequências, obtidos com 13 marcadores SSR, em dez populações de coqueiro-gigante-do-Brasil. Aracaju, SE, 2013.

Loco	Alelo (pb)	Frequência	População
CnCir E2	159	0,017	3 - Baía Formosa
CnCir B12	169	0,036	4 - Georgino Avelino
CnCir C12	171	0,029	4 - Georgino Avelino
CnCir G11	190	0,050	4 - Georgino Avelino
CnCir E2	151	0,042	6 - Santa Rita
CnCir B12	179	0,033	8 - Japoatã
CnCir C12	163	0,077	8 - Japoatã
CnCir E2	145	0,033	8 - Japoatã
CnCir A3	240	0,029	9 - Pacatuba
CnCir B12	181	0,038	9 - Pacatuba
CnCir G11	210	0,167	9 - Pacatuba
CnCir E2	165	0,056	10 - Praia do Forte

Dos 12 alelos exclusivos, oito também podem ser considerados como alelos raros localizados, ou seja, que se encontra em uma única população e com frequência inferior a 5%. Foi encontrado também um alelo raro esporádico, ou seja, com frequência inferior a 5%, mas presente em três das dez populações, sendo estas as populações que se localizam ao sul de Natal, RN.

Vale ressaltar que a questão de alelos exclusivos pode estar diretamente relacionada ao tamanho da amostra e que, portanto, a classificação de um determinado alelo como exclusivo é válido para aquela amostra, uma vez que esses alelos presentes em uma dada população podem não ser capturados por uma amostra pequena.

Em coqueiro, uma amostra composta de 50 a 100 plantas por população, coletadas ao acaso pode ser considerada adequada, para a maioria dos casos, em capturar pelo menos uma cópia de cada alelo que ocorre com frequência maior que 0,05 (MARSHALL; BROWN, 1975). Estes autores sugerem que alelos raros têm provavelmente baixo valor adaptativo e são de menor interesse para os melhoristas. Entretanto, seria estratégico coletar pelo menos uma cópia de cada alelo que ocorre com frequência menor que 0,05. As populações de Santo Inácio, Luís Correia, São José do Mipibu e Merepe não apresentaram alelos exclusivos.

Vale salientar que a questão da presença de alelos exclusivos e raros detectados no presente estudo pode estar relacionada com o tamanho da amostra e que deve ser visto com cuidado, pois esses resultados são válidos para as amostras avaliadas e que, portanto, podem diferir com avaliação de amostras maiores.

A Tabela 6 mostra a classificação dos indivíduos avaliados em dez populações de coqueiro-gigante-do-Brasil, de acordo com a população de referência. Esta classificação se baseia nos escores de cada indivíduo e no inverso do logaritmo na base dez da verossimilhança, classificando-os de acordo com a probabilidade de pertencer ou não à sua população de origem. Foram bem classificados 128 (66%) dos 195 indivíduos analisados, ou seja, aqueles indivíduos cuja probabilidade de pertencer à população de referência é maior do que para outras populações. Pode-se verificar que mais da metade dos indivíduos que não se classificaram em sua população de origem (36 de 67) se encontram em dois grupos de populações. Um primeiro grupo composto pelas populações de Baía Formosa, Georgino Avelino e São José do Mipibu, denominado grupo de Natal e um segundo grupo composto pelas populações de Japoatã, Pacatuba e Praia do Forte, denominado de grupo do Sul. Sete indivíduos da população de Merepe se associaram às populações do grupo de Natal, sendo quatro na população de Georgino Avelino e três na população de Baía Formosa. O número de indivíduos que não foram associados à sua população de origem (67), representa 34% do total dos indivíduos avaliados.

**Tabela 6.** Classificação dos 195 indivíduos avaliados em dez populações de coqueiro-gigante-do-Brasil, de acordo com a população de referência, com base no critério de Rannala e Mountain (1997). Aracaju, SE, 2013.

Origem	1	2	3	4	5	6	7	8	9	10	Total
1 (SI)	<b>14</b>	1				1				1	17
2 (LC)	1	<b>13</b>	1			2					17
3 (BF)	1	1	<b>20</b>	6	1	1					30
4 (GA)			3	<b>13</b>	4					1	21
5 (SJM)			3	3	<b>10</b>			1	1		18
6 (SR)		2			1	<b>16</b>	1			1	21
7 (ME)	1		3	4			<b>11</b>		1		20
8 (JP)	1		1					<b>11</b>	4	1	18
9 (PC)			1					4	<b>13</b>	1	19
10 (PF)					1			4	2	<b>7</b>	14
											<b>128</b>

Considerando-se cada um dos dois grupos como uma única população, obtiveram-se os dados de classificação dos 195 indivíduos associados de acordo com a população de referência, agrupando-se as populações conforme mostrado, na Tabela 7. De acordo com os resultados fizeram-se dois grupos, o grupo de Natal composto pelas populações de Baía Formosa, Georgino Avelino e São José do Mipibu e o grupo do sul composto pelas populações de Japoatã, Pacatuba e Praia do Forte. Após o agrupamento, o número de indivíduos bem classificados passou de 128 para 161 (83%), dos 195 indivíduos analisados. Somente seis indivíduos foram associados preferencialmente a outra população, dos 51 indivíduos que compõem o grupo do Sul e nove indivíduos dos 69 que pertencem ao grupo de Natal.

**Tabela 7.** Classificação dos 195 indivíduos associados de acordo com a população de referência, após a formação de dois grupos. Aracaju, SE, 2013.

Origem	SI	LC	Natal	SR	ME	Sul	Total
SI	<b>16</b>	1					17
LC	1	<b>13</b>	1	2			17
Natal	1	2	<b>60</b>	2	1	3	69
SR		2	1	<b>16</b>	1	1	21
ME	1		7		<b>11</b>	1	20
Sul	1	1	2	1	1	<b>45</b>	51
							<b>161</b>

Vale ressaltar que somente 38% das probabilidades correspondentes foram maiores que 95% e que, portanto, são consideradas como não ambíguas. Entretanto 83% dos indivíduos foram considerados como bem classificados e estão representados com números em negrito. Esses resultados foram utilizados para uma primeira classificação e formação de dois grupos entre as populações. Após o agrupamento o número de indivíduos não associados à população ou grupo de origem, passou de 67 para 34 dos 195 indivíduos analisados, ou seja, 17% (Tabela 7).

### Estruturação da variabilidade genética

O teste de aderência às proporções de equilíbrio de Hardy-Weinberg é mostrado, na Tabela 8. Dos 130 testes de aderência realizados, apenas 16 apresentaram desvios significativos, isto sugere uma aproximação em relação ao esperado pelas proporções de equilíbrio de Hardy-Weinberg. Apenas a população de Baía Formosa apresentou significância no teste, em quatro dos 13 locos. As demais populações apresentaram significância em um ou dois locos dos 13 utilizados.

**Tabela 8.** Probabilidade do teste de aderência ao equilíbrio de Hardy-Weinberg nos 13 locos de SSR, para as dez populações de coqueiro gigante do Brasil. Os valores em negrito indicam os locos em desequilíbrio nas populações. Aracaju, SE, 2013.

Loco	Pop.1	Pop.2	Pop.3	Pop.4	Pop.5	Pop.6	Pop.7	Pop.8	Pop.9	Pop.10
CnCir A3	-	-	-	-	-	-	-	1,000	1,000	1,000
CnCir A9	1,000	0,090	0,001	0,481	0,606	0,655	0,167	0,751	0,931	0,343
CnCir B6	0,642	0,405	0,456	0,652	0,372	0,266	0,009	0,001	0,243	0,519
CnCir B12	0,065	0,970	0,311	0,219	1,000	1,000	0,133	0,718	0,442	0,057
CnCir C7	0,786	1,000	0,395	0,028	0,256	0,290	0,366	0,641	0,630	0,574
CnCir C12	1,000	0,037	0,594	0,158	0,191	0,549	0,487	0,527	0,722	1,000
CnCir E2	0,067	0,724	0,049	0,350	0,386	0,030	0,304	0,462	0,572	0,268
CnCir E10	0,044	0,004	0,045	0,001	1,000	0,524	1,000	0,803	0,451	1,000
CnCir E12	0,138	1,000	1,000	-	0,257	1,000	-	1,000	0,035	1,000
CnCir F2	0,595	0,311	0,025	0,236	1,000	1,000	1,000	0,898	0,228	1,000
CnCir G11	0,412	0,116	0,397	1,000	0,011	0,368	1,000	0,408	0,957	0,491
CnCir H4'	0,053	0,644	1,000	-	1,000	0,653	1,000	0,626	1,000	0,433
CnCir H7	1,000	0,216	0,416	0,067	0,656	0,795	1,000	1,000	0,008	0,029



As estimativas das estatísticas  $F$  de Wright, com os respectivos intervalos de confiança a 95% de probabilidade são apresentadas, na Tabela 9. As estimativas das estatísticas  $F$  de Wright, conforme Weir e Cockerham (1984) foram:  $F = 0,196$ , com intervalo de confiança de 0,151 e 0,234;  $\Phi_p = 0,1600$ , com intervalo de confiança de 0,122 e 0,199; e o  $f = 0,043$ , com intervalo de confiança de 0,011 a 0,073. Portanto estes parâmetros diferem significativamente de zero.

**Tabela 9.** Estimativa das estatísticas  $F$  de Wright, conforme Weir e Cockerham (1984), em 10 populações de coqueiro-gigante-do-Brasil. Intervalo de Confiança (IC) de 95% de probabilidade. Aracaju, SE, 2013.

	$f$	$\Phi_p$	$F$
Estimativa	0,0426	0,1600	0,1958
Limite inferior (IC 95%)	0,0108	0,1222	0,1510
Limite superior (IC 95%)	0,0731	0,1991	0,2345

O parâmetro  $f$  é a correlação de alelos dentro de indivíduos dentro de uma população. Refere-se ao coeficiente de endogamia de um indivíduo devido a cruzamentos não aleatórios dentro de populações, isto é, mede os efeitos da ausência de panmixia dentro das populações.

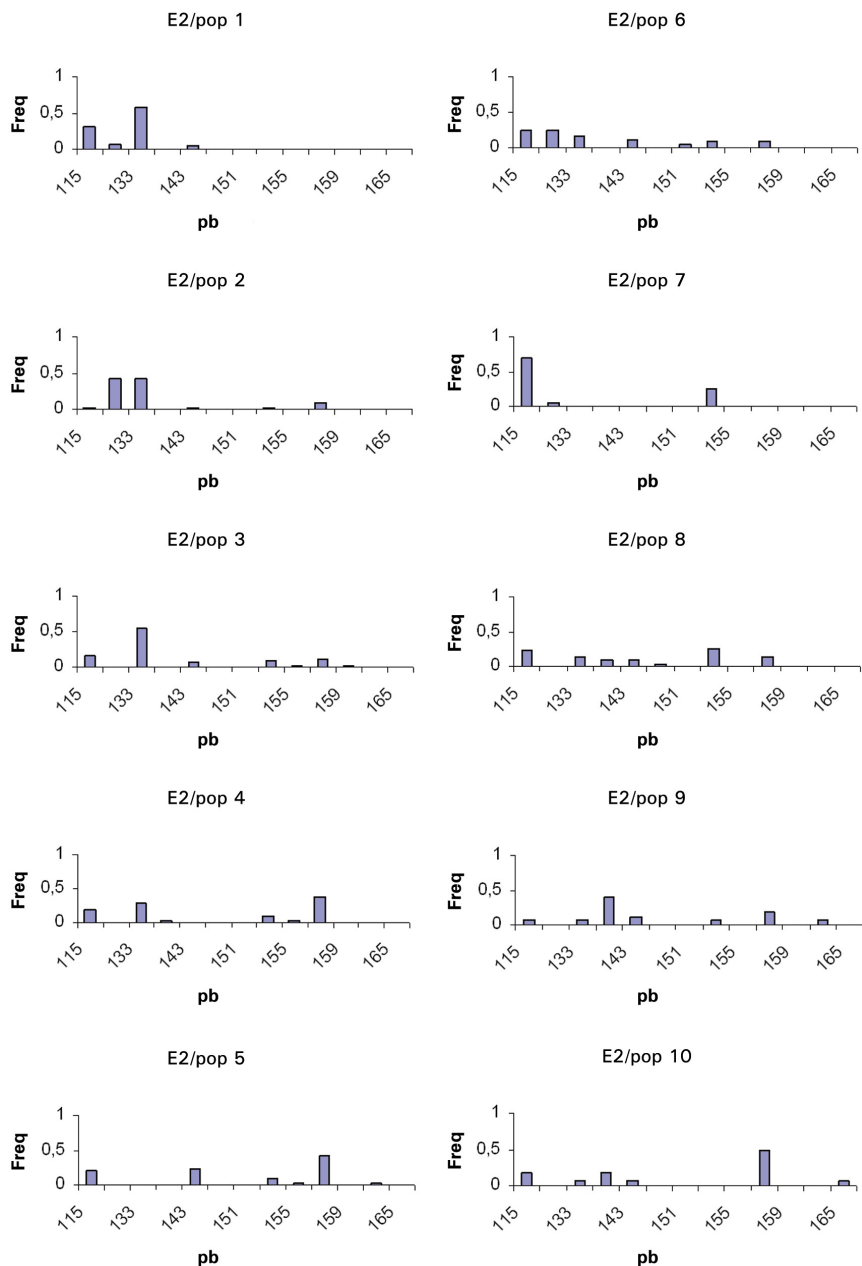
O parâmetro  $\Phi_p$  é a correlação de alelos de diferentes indivíduos numa mesma população, tendo como referência o conjunto das populações. Refere-se ao coeficiente de endogamia dentro das populações, provocado pela deriva genética associada à subdivisão das populações.

Já a estatística  $F$  pode ser entendida como a correlação entre alelos dentro de indivíduos. Refere-se ao coeficiente de endogamia considerando o conjunto das populações, ou seja, é a endogamia total da espécie e resulta dos efeitos da endogamia devida à subdivisão das populações ( $\Phi_p$ ) e da endogamia devida ao sistema reprodutivo ( $f$ ), ou seja, mede os efeitos combinados da ausência de panmixia e da deriva genética.

O índice de fixação intrapopulacional ( $f$ ) foi significativamente diferente de zero, mas com valor de baixa magnitude (0,043) sugerindo não haver grandes desvios de panmixia. Vale ressaltar que a endogamia encontrada dentro das populações é decorrente da geração anterior, ou seja, da geração que originou as sementes, por meio de possíveis cruzamentos entre indivíduos aparentados ou autofecundações.

O valor de  $F_p$  médio foi de 0,160, o qual é significativamente diferente de zero, reflete um alto nível de diferenciação populacional. Este valor indica que 16% da variação total se encontra entre populações e que os restantes 84%, se concentram dentro das populações. Este valor de  $F_p = 0,160$ , indica que a subdivisão, ou seja, a deriva genética tem maior contribuição para a endogamia da espécie do que o sistema reprodutivo.

Embora o modelo de mutação apropriado para marcadores microssatélites seja o de *stepwise*, no presente trabalho a distribuição dos alelos nas populações não se comportaram como esperado, não se ajustando perfeitamente a esse modelo. Na Figura 2, é apresentado um exemplo de distribuição dos alelos nas dez populações de coqueiro-gigante-do-Brasil, de acordo com o tamanho (pb) e frequências dos alelos para o loco CnCir E2, os demais gráficos são apresentados na seção apêndices. Como podem ser observados, os alelos estão distribuídos nas populações de forma aleatória e muitas vezes com frequências maiores nos extremos. Talvez por essas populações apresentarem uma história evolutiva recente e não tenha havido tempo suficiente para que as mutações ocorram dentro das frequências esperadas para o modelo.



**Figura 2.** Exemplo de distribuição alélica em dez populações de coqueiro-gigante-do-Brasil para o loco E2.

Com base nos valores estimados de  $F_{ST}$ , é recomendável que na estratégia de amostragem para coleta de germoplasma seja amostrado um grande número de populações, pela grande variação observada entre as populações.

Esses resultados aliados a dados históricos podem ser considerados como indicio que essas populações se encontram em processo de diferenciação recente, ou seja, ocorreram poucas gerações no processo evolutivo, a partir das populações ancestrais que originaram as populações atuais. Como o coqueiro chegou ao Brasil em 1553, e uma geração dura de 80 a 100 anos, isto implica que ocorreram poucas gerações desses materiais aqui no Brasil e, provavelmente, a estruturação foi influenciada por forte efeito fundador e pela seleção praticada pelo homem.

## Sistema reprodutivo

As estimativas do índice de fixação intrapopulacional ( $f$ ), bem como os intervalos de confiança obtidos por amostragem de 10.000 *bootstraps* são apresentadas, na Tabela 10. Os valores variaram de - 0,100 na população de Pacatuba a 0,134 na população de São José do Mipibu, com média geral de 0,043, conforme já mostrado. Verifica-se que os valores obtidos não foram significativamente diferentes de zero, exceto para a população de Pacatuba onde houve excesso de heterozigotos ( $f < 0$ ).

**Tabela 10.** Dados do índice de fixação intrapopulacional ( $f$ ) após amostragem com reposição dos indivíduos com todos os locos, para as dez populações de coqueiro-gigante-do-Brasil, com amostragem de 10.000 Bootstraps e  $t_a$ : taxa aparente de cruzamento (IC 95%). Aracaju, SE, 2013.

População	$f$	Lim. Inf.	Lim. Sup.	$t_a$
1	- 0,06577	- 0,29560	0,07134	1,141
2	0,01819	- 0,16604	0,12592	0,965
3	0,09898	- 0,02575	0,19301	0,820
4	0,12301	- 0,04205	0,21392	0,781
5	0,13367	- 0,04707	0,25312	0,764
6	- 0,05849	- 0,21674	0,02021	1,123
7	0,10948	- 0,09987	0,25301	0,803
8	0,05944	- 0,11711	0,18039	0,888
9	- 0,09966*	- 0,25122	- 0,03124	1,222
10	0,08470	- 0,08320	0,16329	0,843
<b>Média</b>	<b>0,0426</b>	<b>0,0108</b>	<b>0,0731</b>	<b>0,918</b>

\* Significativo a 5% de probabilidade.

As estimativas das taxas aparentes de cruzamento ( $t_a$ ), obtidas com base neste índice são também apresentadas, na Tabela 10. Os valores desta variaram entre 0,764 na população de São José do Mipibu e 1,222 na população de Pacatuba, com valor médio de 0,918. Isto sugere que a espécie se comporta como preferencialmente alógama.

Estes resultados, associados ao teste de aderência ao equilíbrio de Hardy-Weinberg, mostram que as populações são preferencialmente alógamas, indicando não haver grandes desvios de panmixia. Os valores de  $f$ , por não serem significativamente diferente de zero, sugerem que a maioria dos locos dentro das populações está dentro das proporções esperadas de Hardy-Weinberg, e que no conjunto das populações houve significância para este parâmetro somente em uma população e no caso, com excesso de heterozigotos.

A não significância dos valores do índice de fixação intrapopulacional ( $f$ ) por populações implica que as respectivas taxas aparentes de cruzamento não diferem de 1,0, em cada uma delas. Por outro lado, o valor global 0,92, difere significativamente de 1,0 o que implica em uma taxa aparente de autofecundação de 0,08 ou 8%. Parte desses homozigotos pode se dever a cruzamentos entre indivíduos aparentados na geração anterior.

Não foram encontrados na literatura estimativas de taxa de cruzamento, em coqueiro, a partir de dados de marcadores moleculares. No entanto, os dados deste trabalho confirmam os dados de outros estudos genéticos que afirmam ser o coqueiro-gigante, uma variedade preferencialmente alógama (FRÉMOND et al., 1966).

## Padrão espacial da variabilidade genética

Os valores de  $F_p$  entre pares de populações podem ser observados, na Tabela 11. Os menores valores observados foram entre as populações que compõem os grupos do Sul, Japoatã (8), Pacatuba (9) e Praia do Forte (10), com valores variando de 0,0240 a 0,0497, e entre as populações do grupo de Natal, Baía Formosa, Georgino Avelino e São José do Mipibu, com valores que variaram de 0,0219 a 0,0625. Os maiores valores de  $F_p$  foram observados entre os pares de populações Santo Inácio e Georgino Avelino (0,3156), Santo Inácio e São José do Mipibu (0,2742), Santo Inácio e Merepe (0,2508) e entre as populações de Luís Correia e São José do Mipibu (0,2487).

**Tabela 11.** Valores de  $F_p$  segundo Weir e Cockerham (1984) para os pares de populações, baseado em 13 locos de SSR em dez populações de coqueiro-gigante-do-Brasil. Aracaju, SE, 2013.

POP.	1	2	3	4	5	6	7	8	9
1	*								
2	0,1676	*							
3	0,2391	0,2290	*						
4	<b>0,3156</b>	<b>0,2487</b>	0,0287	*					
5	<b>0,2742</b>	0,2124	0,0625	0,0219	*				
6	0,1030	0,0762	0,1245	0,1621	0,1451	*			
7	<b>0,2508</b>	0,1660	0,1583	0,1498	0,1321	0,1494	*		
8	0,1323	0,1767	0,1382	0,1808	0,1258	0,093	0,1706	*	
9	0,1831	0,2256	0,1767	0,2196	0,1525	0,1446	0,2407	0,0240	*
10	0,1353	0,1959	0,1959	0,2318	0,1544	0,1044	0,2369	0,0425	0,0497

A correlação entre as matrizes de distâncias genética e geográfica entre as dez populações (Tabela 12), foi de  $r = 0.598$ , sendo significativa ( $p = 0.0027$ ) pelo teste de Mantel. Esta correlação foi de intermediária a alta, sinalizando que a divergência genética está estruturada no espaço provavelmente em um padrão clinal de variação. Essas análises indicam que provavelmente um processo

estocástico foi o responsável pela diferenciação genética dessas populações, havendo fluxo a curtas distâncias e deriva genética dentro das populações locais.

**Tabela 12.** Matriz das distâncias genéticas de Nei (acima da diagonal) e distâncias geográficas (abaixo da diagonal) entre dez populações de coqueiro-gigante-do-Brasil. Aracaju, SE, 2013.

Pop.	1	2	3	4	5	6	7	8	9	10
1	0	0.209	0.294	0.390	0.381	0.141	0.318	0.197	0.258	0.197
2	135,637	0	0.275	0.272	0.267	0.107	0.188	0.272	0.335	0.294
3	952,959	818,839	0	0.034	0.071	0.145	0.163	0.174	0.214	0.255
4	938,864	804,351	23,2024	0	0.035	0.171	0.137	0.214	0.251	0.279
5	921,216	786,701	36,7638	17,651	0	0.189	0.145	0.178	0.198	0.213
6	1051,245	921,196	163,350	186,481	196,790	0	0.182	0.153	0.219	0.171
7	1081,458	954,122	232,119	255,002	263,929	70,989	0	0.237	0.335	0.343
8	1101,406	992,039	487,137	506,586	508,564	353,226	290,281	0	0.063	0.099
9	1114,995	1005,137	489,055	508,868	511,350	351,886	287,497	14,834	0	0.096
10	1227,936	1138,213	759,819	778,455	779,186	627,800	563,180	274,915	275,980	0

As matrizes de distâncias genéticas e geográficas entre as dez populações são apresentadas, na Tabela 12. As distâncias genéticas obtidas entre as populações variaram de 0,034 a 0,390. As menores distâncias genéticas observadas foram entre as populações de Baía Formosa e Georgino Avelino, Georgino Avelino e São José do Mipibu (5), 0,034 e 0,035, respectivamente e entre Georgino Avelino e São José do Mipibu, 0,071. Também apresentaram pequenas distâncias genéticas entre si as populações de Japoatã e Pacatuba (0,063), Pacatuba e Praia do Forte (0,096) e Japoatã e Praia do Forte (0,099). Estes menores valores coincidem com aqueles encontrados para  $D_p$  par a par.

Os dados da Tabela 12 sugerem a possibilidade de agrupamento das populações com distâncias genéticas menores que 0,1. Com isso, formaram-se dois grupos de populações, o grupo 1 composto pelas populações de Baía Formosa, Georgino Avelino e São José do Mipibu, denominado grupo de Natal, e o grupo 2 formado pelas populações de Japoatã, Pacatuba e Praia do Forte, denominado grupo do Sul. Pode se verificar que as populações mais similares geneticamente são também as populações geograficamente mais próximas. As demais populações apresentam comportamento diferenciado sem possibilidade de formação de grupos consistentes. Este agrupamento corrobora com aquele proposto com base na classificação pelo critério de Ranalla e Mountain (1997).

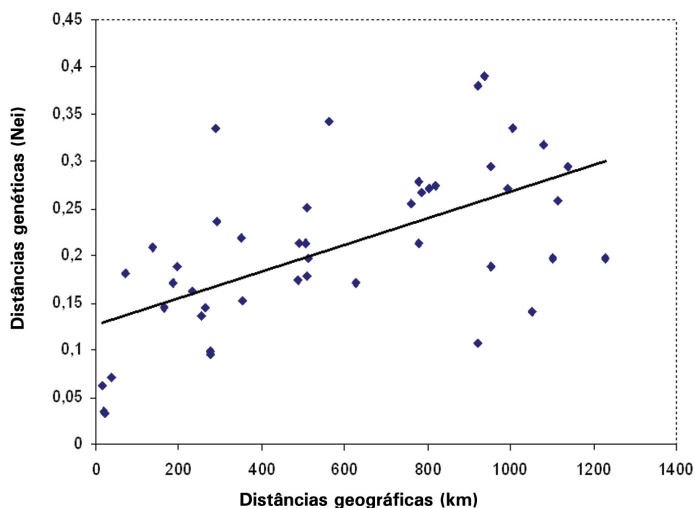
As maiores distâncias genéticas observadas foram entre as populações de Santo Inácio e Georgino Avelino (0,390), Santo Inácio e São José do Mipibu (0,381), Merepe e Praia do Forte (0,343), Merepe e Pacatuba (0,335) e Luís Correia e Pacatuba (0,335), as quais também apresentam grandes distâncias geográficas (Tabela 12).

O padrão da correlação entre as matrizes de distâncias genéticas de Nei e as distâncias geográficas entre as dez populações de coqueiro-gigante-do-Brasil é apresentado na Figura 3. A correlação matricial entre as distâncias genéticas e as distâncias geográficas foi positiva ( $r = 0,598$ ) e significativa a 1% de probabilidade ( $p = 0,0027$ ) pelo teste de Mantel com 10.000 permutações. Este resultado sugere um padrão espacial da variabilidade genética entre as populações, e que as mesmas estão estruturadas no espaço.



$$r = 0,598$$

$$P = 0,0027 \text{ (10000 permutações)}$$

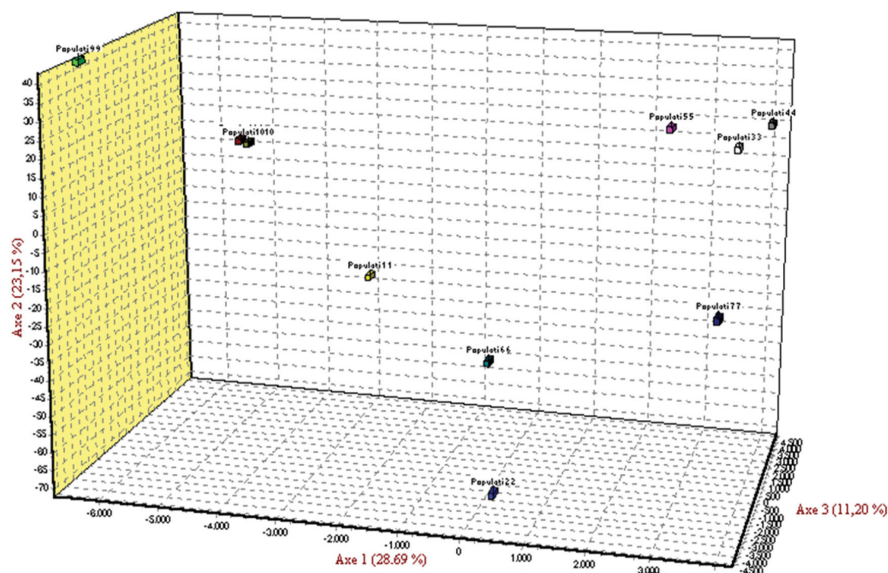


**Figura 3.** Representação da correlação entre as distâncias genéticas e as distâncias geográficas entre dez populações de coqueiro-gigante-do-Brasil. Aracaju, SE, 2013.

Pela Figura 3, constata-se que populações geograficamente mais próximas possuem consistentemente menores distâncias genéticas. Para populações geograficamente distantes ocorrem altos e baixos valores para a distância genética, em um padrão mais aleatório. Este padrão é coerente com o modelo de isolamento pela distância em populações naturais (Wright, 1943).

## Aspectos histórico-evolutivos

O padrão de distribuição das dez populações de coqueiro-gigante-do-Brasil pode ser visualizado no gráfico tridimensional apresentado, na Figura 4. A análise fatorial de correspondência (AFC) permite visualizar mais facilmente as posições relativas das amostras. A visualização pode ser efetuada no espaço tridimensional, no qual cada eixo representa a sua contribuição para a variação total. A contribuição relativa dos indivíduos, para a inércia (variação) dos três primeiros eixos da AFC permite descrever a variabilidade sobre os três planos e acumulam 63% da variação total. As inércias (variações) relativas dos três primeiros eixos decrescem de 28,69% a 11,20%.

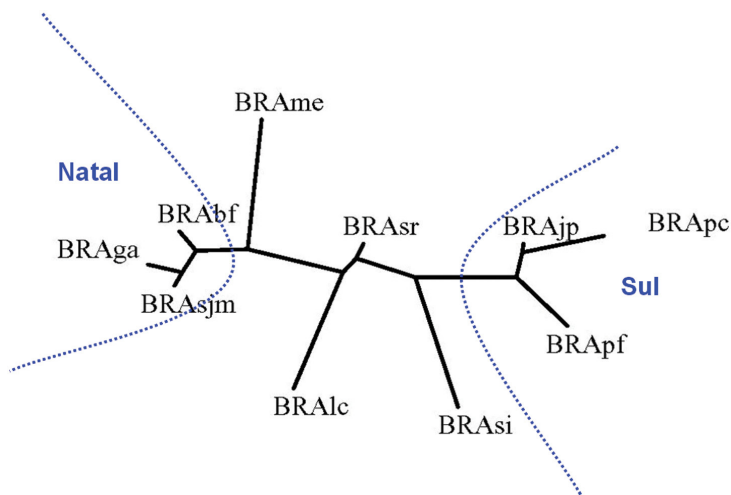


**Figura 4.** Gráfico tridimensional representando o padrão de distribuição das dez populações de coqueiro-gigante-do-Brasil, obtido pela análise fatorial de correspondência (AFC). Aracaju, SE, 2013.

Com base nos resultados da AFC pode-se identificar no canto superior direito do gráfico um conjunto de populações que corresponde as populações Baía Formosa (3), Georgino Avelino (4) e São José do Mipibu (5). Este primeiro grupo foi denominado grupo de Natal. Um outro conjunto de populações pode ser visto na parte superior esquerda e corresponde as populações Japoatã (8), Pacatuba (9) e Praia do Forte (10). Esse segundo grupo foi denominado de grupo do Sul. Esses dois grupos são formados por três populações cada. As demais populações se encontram dispersas no espaço multidimensional e não apresentaram nenhuma tendência de agrupamento. Estes dados corroboram com o agrupamento proposto com base na matriz de distâncias genéticas (Tabela 12).

O padrão de divergência genética entre as populações de coqueiro-gigante-do-Brasil foi avaliado também com base no dendrograma, apresentado na Figura 5, obtido a partir das distâncias genéticas (NEI, 1972). Pela análise dos dados pode-se observar o padrão de divergência entre as dez populações confirmando-se que as populações 3 (Baía Formosa), 4 (Georgino Avelino) e 5 (São José do Mipibu) foram geneticamente semelhantes entre si e formando um grupo diferenciado das

demais populações. Essas populações se localizam próximo à cidade de Natal, RN. Fato semelhante observa-se com as populações 8 (Japoatã), 9 (Pacatuba) e 10 (Praia do Forte) que também apresentaram similaridade genética entre si, formando um segundo grupo denominado grupo do Sul, pois as populações se localizam no Sul da região de ocorrência das populações compreendida entre o norte de Sergipe e norte da Bahia. As demais populações apresentaram padrões variados de divergência e se comportaram de forma diferente.



**Figura 5.** Padrão de divergência genética entre dez populações de coqueiro-gigante-do-Brasil, obtido pelo método de Neighbor Joining, com base nas distâncias genéticas Aracaju, SE, 2013.

O mapa da Figura 6 mostra a posição geográfica das dez populações com os grupos formados. Cada grupo está representado por um círculo azul que envolve as populações de acordo com o agrupamento efetuado.



**Figura 6.** Mapa de localização das dez populações de coqueiro-gigante-do-Brasil, de acordo com o agrupamento, utilizando 13 locos SSR, com base nas distâncias genéticas e na análise fatorial de correspondência (AFC). Aracaju, SE, 2013.

A diversidade dos coqueiros gigantes é dividida em duas grandes zonas: o grupo que corresponde à zona ecogeográfica do Pacífico, compreende o Sudeste asiático e o Pacífico Sul e é denominado “grupo Pacífico”, e o segundo grupo que corresponde aos coqueiros da Índia, Sri Lanka e do Oeste da África, este grupo é denominado “grupo Indo-Atlântico” (BOURDEIX et al., 1997; LEBRUN et al., 1999), o qual as populações brasileiras pertencem.

A dispersão do coqueiro, a partir do Oeste da África (Cabo Verde), para o caribe e Costa Atlântica da América, aliada à semelhança morfológica entre esses materiais e os coqueiros brasileiros, são fortes argumentos que permitem

considerar que eles pertencem ao mesmo grupo, ou seja, ao grupo Indo-Atlântico (LEBRUN et al., 1999).

## Considerações Finais

As populações brasileiras apesar de terem uma história evolutiva comum e recente apresentam grande divergência genética, provavelmente devido ao seu modo de dispersão associado ao homem, que muitas vezes pratica seleção de poucos indivíduos para o estabelecimento de novas populações e isto pode ter levado a um alto grau de diferenciação entre elas.

Com base nos resultados obtidos, e devido à natureza recalcitrante das sementes, podem se adotar estratégias mais eficientes de coleta e conservação do germoplasma *ex situ* de coco. Para a coleta pode se constituir uma amostra para cada grupo (Natal e do Sul), composta de uma mistura de sementes das populações dentro de cada grupo, para representar um acesso por grupo. Em relação às demais populações, poder-se-ia se coletar uma amostra por população, constituída por um grande número de indivíduos coletados, para capturar a maioria dos alelos de interesse presentes nessas populações.

Embora este trabalho tenha fornecido importantes resultados, é importante incluir nos próximos estudos os materiais de coqueiros exóticos, principalmente os materiais de origem asiática, e materiais da variedade anã para que se tenham informações mais completas sobre o germoplasma de coqueiro brasileiro e com isso tornar mais eficiente o programa de melhoramento da cultura.

Diante das observações de uma alta porcentagem de alelos comuns dispersos (82,35%) aliado ao alto nível de variação dentro das populações (84%), sugere-se que coletar uma amostra grande aleatória das populações alvo, seria suficiente para capturar a maioria da variação genética desejável presente nestas populações. No caso das populações que se agrupam, uma grande amostra por grupo seria o suficiente para esse propósito.

A grande estruturação genética, medida pela estatística  $F_{st}$ , reforça a necessidade de que todas as populações sejam amostradas para conservação *ex situ*.

## Conclusões

- As populações de coqueiro-gigante-do-Brasil apresentam alta estruturação genética detectada pelos marcadores SSR, com 16% da variabilidade genética entre e 84% dentro das populações.
- A variabilidade genética nas populações estudadas está espacialmente estruturada, com maior similaridade entre populações geograficamente mais próximas.
- As populações Baía Formosa, Georgino Avelino e São José do Mipibu formam um grupo de similaridade denominado grupo de Natal e as populações Japoatã, Pacatuba e Praia do forte formam um segundo grupo de similaridade denominado grupo do Sul.
- As populações de coqueiro-gigante-do-Brasil são preferencialmente alógamas, com taxa média aparente de cruzamento de 92%.
- As populações brasileiras possuem origem única e apresentam alta divergência entre si, em comparação com outras populações do grupo Indo-Atlântico.

## Referências

BAUDOUIN, L.; LEBRUN, P. The development of a microsatellite kit and dedicated software for use with coconuts. Rome: International Plant Genetic Resources Institute (IPGRI). **Burotrop Bulletin**, Paris, v. 17, p. 16-20, 2002.

BELKHIR, K.; BORSA, P.; CHIKHI, L.; RAUFASTE, N.; BONHOMME, F. **Genetix 4.03. Logiciel sous windows TM pour la génétique des populations**. Laboratoire génome, populations, interactions, CNRS UMR 5000, Université Montpellier II, Montpellier, France, 2001.

BENOIT, H.; GHESQUIÈRE, M. **Electrophorèse, compte rendu cocotier. IV. Déterminisme génétique**. Montpellier: IRHO-CIRAD, 11p. 1984. Rapport interne.

BOURDEIX, R.; BAUDOUIN, L.; BILLOTE, N.; LABOUISSÉ, J. P.; NOIRET, J. M. Le cocotier. In: CHARRIER, A.; JACQUOT, M.; HAMON, S.; NICOLAS, D. (Ed.). **L'amélioration des plantes tropicales**. Montpellier: CIRAD-ORSTOM, 1997. p. 217-239.

COCKERHAM, C. C. Variance of gene frequencies. **Evolution**. Lancaster, v. 23, p. 72-84, 1969.

COCKERHAM, C. C. Analyses of gene frequencies. **Genetics**, Bethesda, v. 74, p. 679-700, 1973.

CHASE, M.; KESSELI, R.; BAWA, K. Microsatellite markers for population and conservation genetics of tropical trees. **American Journal of Botany**, Bronx, v. 83, n. 1, p. 51-57, 1996.

FAO. (Roma). World Production. Disponível em: site **FAO** URL: <<http://apps.fao.org/page/collection?subset=agriculture>>. Consultado em 23 de Novembro de 2005.

FELSENSTEIN, J. **PHYLIP (Phylogeny Inference Package)**, version 3.6. Seattle: Department of Genome Sciences, University of Washington, 2004.

FRÉMOND, Y.; ZILLER, R.; NUCÉ de LAMOTHE, M. de. **Le cocotier**. Paris: Maison-Neuve & Larose, 1966. 267 p.

JAY, M.; BOURDEIX R.; POTIER, F.; SANLAVILLE, C. Premier résultats de l'étude des polyphénols foliaires du cocotier. **Oléagineux**, Paris, v. 44, p. 151-161, 1989.

LEBRUN, P.; BAUDOUIN, L.; SÉGUIN, M.; N'CHO, Y. P.; BOURDEIX, R. Etude de la diversité génétique du cocotier par RFLP. **OCL**, Paris, v. 2, n.6, p. 418-420, 1995.

LEBRUN, P.; N'CHO, Y. P.; BOURDEIX, R.; BAUDOUIN, L. Le cocotier. In: HAMON, P.; SEGUIN, M.; PERRIER, X.; GLASZMAN, J. C. (Ed.). **Diversité génétique des plantes tropicales cultivées**. Montpellier: CIRAD, 1999. p. 219-239.

MANLY, B. F. J. **Multivariate statistical methods: a primer**, London: Chapman & Hall, 1986a, 159 p.

MANLY, B.F.J. Randomization and regression methods for testing for associations with geographical, environmental and biological distances between populations. **Research Population Ecology**, Tokyo, v. 28, p. 201-208, 1986b.

MANTEL, N. The detection of disease clustering and a generalized regression approach. **Cancer Research**, Baltimore, US, v. 27, p. 209-220, 1967.

MARSHALL, D. K.; BROWN, A. H. D. Optimum sampling strategies in genetic conservation. In: Frankel, O. H.; Hawkes, J. G. R. (Ed.). **Crop genetic resources for today and tomorrow**, Cambridge: Cambridge University Press, 1975. p. 53-70.

MILLER, M. **Tools for Population Genetic Analyses (TFPGA) 1.3**: A windows program for analyses of allozyme and molecular population genetic data. 1997.



MORGANTE, M.; OLIVIERI, A. M. PCR-amplified of microsatellite as markers in plant genetics. **Plant Journal**, Oxford, v. 3, p. 175-182, 1993.

MORGANTE, M.; PFEIFFER, A.; COSTACURTA, A.; OLIVIERI, A. M. Molecular tools for population and ecological genetics in coniferous trees. **Phyton**, Horn, v. 36, n. 3, p. 129-138, 1996.

NEI, M. Genetics distance between populations. **American Naturalist**, Chicago, v. 106, p. 283-292, 1972.

PERERA, L.; RUSSEL, J. R.; PROVAN, J.; POWEL, W. Use of microsatellite DNA markers to investigate the level of genetic diversity and population genetic structure of coconut (*Cocos nucifera* L.). **Genome**, Ottawa, v. 43, p. 15-21, 2000.

PERERA, L.; RUSSEL, J. R.; PROVAN, J.; POWEL, W. Level and distribution of genetic diversity of coconut (*Cocos nucifera* L. var. *Typica* form typical ) from Sri Lanka assessed by micorsatellite markers. **Euphytica**, Wageningen, v. 122, p. 381-389, 2001.

PERRIER, X.; FLORI, A. ; BONNOT, F. Les méthodes d'analyse des données. IN: HAMON, P.; SEGUIN, M.; PERRIER, X.; GLASZMANN, J. C. **Diversité génétique des plantes tropicales cultivées**. Montpellier: CIRAD, 1999. p. 43-76.

PERSLEY, G. J. **Replanting the tree of life: toward an international agenda for coconut palm research**. Wallingford: CAB/ACIAR, 1992. 156 p.

PIRY, S.; ALAPETITE, A.; CORNUET, J.M.; PAETKAU, D.; BAUDOUIN, L.; ESTOUP, A. GeneClass2: a software for genetic assignment and first generation migrants detection. **Journal of Heredity**, Washington, v. 95, p. 536-539, 2004.

PURSEGLOVE, J. W. **Tropical crops: monocotyledons**. Burnt Will: Longman, 1975. 607 p.

RANNALA, B.; MOUNTAIN, J. L. Detecting immigration by using multilocus genotypes. **Proceedings of the National Academy of Sciences**, Washington, v. 94, p. 9197-9201, 1997.

RIBEIRO, F. E. **Divergência genética entre populações de coqueiro gigante (*Cocos nucifera* L.) do Brasil**. 1993. 84 f. Dissertação (Mestrado) - Escola Superior de Agricultura de Lavras, Lavras, 1993.

RIBEIRO, F. E.; SOARES, A. R.; RAMALHO, M. A. P. Caracterização física dos frutos de cinco populações de coqueiro gigante (*Cocos nucifera* L.) no Nordeste do Brasil. **Revista Científica Rural**, Bagé, v. 2, n. 1, p. 28-34, 1997.

RIBEIRO, F. E.; SOARES, A. R.; RAMALHO, M. A. P. Divergência genética entre populações de coqueiro gigante do Brasil. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 34, n. 9, p. 1615-1622, 1999.

RIBEIRO, F. E.; SIQUEIRA, E. R. de; ARAGÃO, W. M.; TUPINAMBÁ, E. A. **Ecótipos de coqueiro gigante no Brasil**. Aracaju: EMBRAPA-CPATC, 2000. 25 p. (EMBRAPA-CPATC. Documentos, 17).

RIBEIRO, F. E.; SIQUEIRA, E. R. de.; ARAGÃO, W. M. Coqueiro. In: BRUCKNER, C. H. (Ed.). **Melhoramentos de fruteiras tropicais**. Viçosa, MG: Universidade Federal de Viçosa, 2002. p. 225-249.

RIVERA, R.; EDWARDS, K. J.; BARKER, J. H. A.; ARNOLD, G. M.; AYAD, G.; HODGKIM, T. Isolation and characterization of polymorphic microsatellites in *Cocos nucifera* L. **Genome**, Ottawa, v. 42, p. 668-675, 1999.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology Evolution**, Chicago, US, v. 4, p. 406-425, 1987.

SIQUEIRA, E. R.; RIBEIRO, F. E.; ARAGÃO, W. M. Melhoramento genético do coqueiro. In: FERREIRA, J. M. S.; WARWICK, D. R. N.; SIQUEIRA, L. A. (Ed.). **A cultura do coqueiro no Brasil**. Aracaju, EMBRAPA-CPATC, 1994. p. 87-120.

VENCOVSKY, R. Análise de variância de frequências alélicas. **Genetics and Molecular Biology**, Ribeirão Preto, v. 15, p. 53-60, 1992.

VENCOVSKY, R. Variance of an estimative of the outcrossing rate. **Brazilian Journal of Genetics**, Ribeirão Preto, v. 17, n. 3, p. 349-351, 1994.

WADT, L. H. de O.; SAKIYAMA, N. S.; FERREIRA, M. G.; TUPINAMBÁ, E. A.; RIBEIRO, F. E.; ARAGÃO, W. M. Divergência genética entre coqueiros gigante avaliada por RAPD com amostras individuais e compostas. **Revista Brasileira de Genética**, Ribeirão Preto, v. 20, n. 3, p. 262, 1997. Suplemento, ref. G.76. Edição dos Resumos do 43º Congresso Nacional de Genética, Goiânia, ago. 1997.

WEIR, B. S.; COCKERHAM, C. C. Estimating F-statistics for the analysis of population structure. **Evolution**, Lancaster, v. 38, p. 1358-1370, 1984.

WEIR B. S. **Genetics data analysis II**: methods for discrete population genetic data. Sunderland: Sinauer Associates, 1996. 455 p.

WRIGHT, S. Isolation by distance. **Genetics**, Bethesda, v. 28, p. 114-138, 1943.

WRIGHT, S. The genetical structure of populations. **Annals of Eugenics**, London, v. 15, p. 323-354, 1951.



---

*Tabuleiros Costeiros*

Ministério da  
**Agricultura, Pecuária  
e Abastecimento**

